OSCO

5 OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/023,888

DATE: 01/15/2002
TIME: 19:00:25

Input Set : A:\203515US77.txt
Output Set: N:\CRF3\01152002\J023888.raw

```
 3 <110> APPLICANT: CANFIELD, William
 5 <120> TITLE OF INVENTION: SOLUBLE GLcNAc PHOSPHOTRANSFERASE
 7 <130> FILE REFERENCE: 203515US77
 9 <140> CURRENT APPLICATION NUMBER: US/10/023,888
 9 <141> CURRENT FILING DATE: 2001-12-21
 9 <160> NUMBER OF SEQ ID NOS: 38
11 <170> SOFTWARE: PatentIn version 3.1
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 3600
15 <212> TYPE: DNA
16 <213> ORGANISM: hybrid
18 <400> SEQUENCE: 1
19 atggagacag acacactcct gctatgggta ctgctgctct gggttccagg ttccactggt    60
21 gacgaagatc aggtagatcc gcggttaatc gacggtaagc ttagccgaga tcaataccat   120
23 gtttttgtttg attcctatag agacaatatt gctggaaagt cctttcagaa tcggctttgt   180
25 ctgcccatgc cgattgacgt tgtttacacc tgggtgaatg gcacagatct tgaactactg   240
27 aaggaactac agcaggtcag agaacagatg gaggaggagc agaaagcaat gagagaaatc   300
29 cttgggaaaa acacaacgga acctactaag aagagtgaga agcagttaga gtgtttgcta   360
31 acacactgca ttaaggtgcc aatgcttgtc ctggacccag ccctgccagc caacatcacc   420
33 ctgaaggacc tgccatctct ttatccttct tttcattctg ccagtgacat tttcaatgtt   480
35 gcaaaaccaa aaaacccttc taccaatgtc tcagttgttg tttttgacag tactaaggat   540
37 gttgaagatg cccactctgg actgcttaaa ggaaatagca gacagacagt atggagggGc   600
39 tacttgacaa cagataaaga agtccctgga ttagtgctaa tgcaagattt ggcttctctg   660
41 agtggatttc caccaacatt caaggaaaca aatcaactaa aaacaaaatt gccagaaaat   720
43 ctttcctcta aagtcaaact gttgcagttg tattcagagg ccagtgtagc gcttctaaaa   780
45 ctgaataacc ccaaggattt tcaagaattg aataagcaaa ctaagaagaa catgaccatt   840
47 gatggaaaag aactgaccat aagtcctgca tatttattat gggatctgag cgccatcagc   900
49 cagtctaagc aggatgaaga catctctgcc agtcgttttg aagatatttt cattgtcacc   960
51 tactcattgc gatctatcga gaggcatgca ccatgggttc ggaatatttt cattgtcacc  1020
53 aacgggcaga ttccatcctg gctgaacctt gacaatcctc gagtgacaat agtaacacac  1080
55 caggatgttt ttcgaaattt gagccacttg cctaccttta gttcacctgc tattgaaagt  1140
57 cacgttcatc gcatcgaagg gctgtcccag aagtttattt acctaaatga tgatgtcatg  1200
59 tttgggaagg atgtctggcc agatgatttt tacagtcact ccaaaggcca gaaggtttat  1260
61 ttgacatggc ctgtgccaaa ctgtgccgag ggctgcccag gttcctggat taaggatggc  1320
63 tattgtgaca aggcttgtaa taattcagcc tgcgattggg atggtgggga ttgctctgga  1380
65 aacagtggag ggagtcgcta tattgcagga ggtggaggta ctgggagtat tggagttgga  1440
67 cagccctggc agtttggtgg aggaataaac agtgtctctt actgtaatca gggatgtgcg  1500
69 aattcctggc tcgctgataa gttctgtgac caagcatgca atgtcttgtc ctgtgggttt  1560
71 gatgctggcg actgtgggca agatcatttt catgaattgt ataaagtgat ccttctccca  1620
73 aaccagactc actatattat tccaaaaggt gaatgcctgc cttatttcag ctttgcagaa  1680
75 gtagccaaaa gaggagttga aggtgcctat agtgacaatc caataattcg acatgcttct  1740
77 attgccaaca gtggaaaaac catccacctc ataatgcaca gtggaatgaa tgccaccaca  1800
79 atacatttta atctcacgtt tcaaaataca aacgatgaag agttcaaaat gcagataaca  1860
81 gtggaggtgg acacaaggga gggaccaaaa ctgaattcta cggcccagaa gggttacgaa  1920
83 aatttagtta gtcccataac acttcttcca gaggcggaaa tccttttttga ggatattccc  1980
85 aaagaaaaac gcttcccgaa gtttaagaga catgatgtta actcaacaag gagagcccag  2040
87 gaagaggtga aaattcccct ggtaaatatt tcactccttc caaaagacgc ccagttgagt  2100
```

Input Set : A:\203515US77.txt
Output Set: N:\CRF3\01152002\J023888.raw

```
 89 ctcaatacct tggatttgca actggaacat ggagacatca ctttgaaagg atacaatttg   2160
 91 tccaagtcag ccttgctgag atcatttctg atgaactcac agcatgctaa aataaaaaat   2220
 93 caagctataa taacagatga aacaaatgac agtttggtgg ctccacagga aaaacaggtt   2280
 95 cataaaagca tcttgccaaa cagcttagga gtgtctgaaa gattgcagag gttgactttt   2340
 97 cctgcagtga gtgtaaaagt gaatggtcat gaccagggtc agaatccacc cctggacttg   2400
 99 gagaccacag caagatttag agtggaaact cacacccaaa aaaccatagg cggaaatgtg   2460
101 acaaaagaaa agcccccatc tctgattgtt ccactggaaa gccagatgac aaaagaaaag   2520
103 aaaatcacag ggaaagaaaa agagaacagt agaatggagg aaaatgctga aaatcacata   2580
105 ggcgttactg aagtgttact tggaagaaag ctgcagcatt acacagatag ttacttgggc   2640
107 tttttgccat gggagaaaaa aaagtatttc ctagatcttc tcgacgaaga agagtcattg   2700
109 aagacacaat tggcctactt cactgatagc aagaatagag ccagatacaa gagagataca   2760
111 tttgcagatt ccctcagata tgtaaataaa attctaaata gcaagtttgg attcacatcg   2820
113 cggaaagtcc ctgctcacat gcctcacatg attgaccgga ttgttatgca agaactgcaa   2880
115 gatatgttcc ctgaagaatt tgacaagacg tcatttcaca aagtgcgcca ttctgaggat   2940
117 atgcagtttg ccttctctta ttttattat ctcatgagtg cagtgcagcc actgaatata   3000
119 tctcaagtct ttgatgaagt tgatacagat caatctggtg tcttgtctga cagagaaatc   3060
121 cgaacactgg ctaccagaat tcacgaactg ccgttaagtt tgcaggattt gacaggtctg   3120
123 gaacacatgc taataaattg ctcaaaaatg cttcctgctg atatcacgca gctaaataat   3180
125 attccaccaa ctcaggaatc ctactatgat cccaacctgc caccggtcac taaaagtcta   3240
127 gtaacaaact gtaaaccagt aactgacaaa atccacaaag catataagga caaaaacaaa   3300
129 tataggtttg aaatcatggg agaagaagaa atcgctttta aaatgattcg taccaacgtt   3360
131 tctcatgtgg ttggccagtt ggatgacata agaaaaaacc ctaggaagtt tgtttgcctg   3420
133 aatgacaaca ttgaccacaa tcataaagat gctcagacag tgaaggctgt tctcagggac   3480
135 ttctatgaat ccatgttccc cataccttcc caatttgaac tgccaagaga gtatcgaaac   3540
137 cgtttccttc atatgcatga gctgcaggaa tggagggctt atcgagacaa attgaagtag   3600
140 <210> SEQ ID NO: 2
141 <211> LENGTH: 1199
142 <212> TYPE: PRT
143 <213> ORGANISM: hybrid
145 <400> SEQUENCE: 2
147 Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
148 1               5                   10                  15
151 Gly Ser Thr Gly Asp Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly
152             20                  25                  30
155 Lys Leu Ser Arg Asp Gln Tyr His Val Leu Phe Asp Ser Tyr Arg Asp
156         35                  40                  45
159 Asn Ile Ala Gly Lys Ser Phe Gln Asn Arg Leu Cys Leu Pro Met Pro
160     50                  55                  60
163 Ile Asp Val Val Tyr Thr Trp Val Asn Gly Thr Asp Leu Glu Leu Leu
164 65                  70                  75                  80
167 Lys Glu Leu Gln Gln Val Arg Glu Gln Met Glu Glu Glu Gln Lys Ala
168                 85                  90                  95
171 Met Arg Glu Ile Leu Gly Lys Asn Thr Thr Glu Pro Thr Lys Lys Ser
172             100                 105                 110
175 Glu Lys Gln Leu Glu Cys Leu Leu Thr His Cys Ile Lys Val Pro Met
176         115                 120                 125
179 Leu Val Leu Asp Pro Ala Leu Pro Ala Asn Ile Thr Leu Lys Asp Leu
180     130                 135                 140
183 Pro Ser Leu Tyr Pro Ser Phe His Ser Ala Ser Asp Ile Phe Asn Val
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/023,888

Input Set : A:\203515US77.txt
Output Set: N:\CRF3\01152002\J023888.raw

```
184 145                        150              155                160
187 Ala Lys Pro Lys Asn Pro Ser Thr Asn Val Ser Val Val Val Phe Asp
188                165                170               175
191 Ser Thr Lys Asp Val Glu Asp Ala His Ser Gly Leu Leu Lys Gly Asn
192            180                185                190
195 Ser Arg Gln Thr Val Trp Arg Gly Tyr Leu Thr Thr Asp Lys Glu Val
196          195                200                205
199 Pro Gly Leu Val Leu Met Gln Asp Leu Ala Phe Leu Ser Gly Phe Pro
200      210                215                220
203 Pro Thr Phe Lys Glu Thr Asn Gln Leu Lys Thr Lys Leu Pro Glu Asn
204 225                230                235                240
207 Leu Ser Ser Lys Val Lys Leu Leu Gln Leu Tyr Ser Glu Ala Ser Val
208                245                250                255
211 Ala Leu Leu Lys Leu Asn Asn Pro Lys Asp Phe Gln Glu Leu Asn Lys
212            260                265                270
215 Gln Thr Lys Lys Asn Met Thr Ile Asp Gly Lys Glu Leu Thr Ile Ser
216          275                280                285
219 Pro Ala Tyr Leu Leu Trp Asp Leu Ser Ala Ile Ser Gln Ser Lys Gln
220      290                295                300
223 Asp Glu Asp Ile Ser Ala Ser Arg Phe Glu Asp Asn Glu Glu Leu Arg
224 305                310                315                320
227 Tyr Ser Leu Arg Ser Ile Glu Arg His Ala Pro Trp Val Arg Asn Ile
228                325                330                335
231 Phe Ile Val Thr Asn Gly Gln Ile Pro Ser Trp Leu Asn Leu Asp Asn
232            340                345                350
235 Pro Arg Val Thr Ile Val Thr His Gln Asp Val Phe Arg Asn Leu Ser
236          355                360                365
239 His Leu Pro Thr Phe Ser Ser Pro Ala Ile Glu Ser His Val His Arg
240      370                375                380
243 Ile Glu Gly Leu Ser Gln Lys Phe Ile Tyr Leu Asn Asp Asp Val Met
244 385                390                395                400
247 Phe Gly Lys Asp Val Trp Pro Asp Asp Phe Tyr Ser His Ser Lys Gly
248                405                410                415
251 Gln Lys Val Tyr Leu Thr Trp Pro Val Pro Asn Cys Ala Glu Gly Cys
252            420                425                430
255 Pro Gly Ser Trp Ile Lys Asp Gly Tyr Cys Asp Lys Ala Cys Asn Asn
256          435                440                445
259 Ser Ala Cys Asp Trp Asp Gly Gly Asp Cys Ser Gly Asn Ser Gly Gly
260      450                455                460
263 Ser Arg Tyr Ile Ala Gly Gly Gly Gly Thr Gly Ser Ile Gly Val Gly
264 465                470                475                480
267 Gln Pro Trp Gln Phe Gly Gly Gly Ile Asn Ser Val Ser Tyr Cys Asn
268                485                490                495
271 Gln Gly Cys Ala Asn Ser Trp Leu Ala Asp Lys Phe Cys Asp Gln Ala
272            500                505                510
275 Cys Asn Val Leu Ser Cys Gly Phe Asp Ala Gly Asp Cys Gly Gln Asp
276          515                520                525
279 His Phe His Glu Leu Tyr Lys Val Ile Leu Leu Pro Asn Gln Thr His
280      530                535                540
```

```
283 Tyr Ile Ile Pro Lys Gly Glu Cys Leu Pro Tyr Phe Ser Phe Ala Glu
284 545               550           555           Phe     560
287 Val Ala Lys Arg Gly Val Glu Gly Ala Tyr Ser Asp Asn Pro Ile Ile
288             565           570                   575
291 Arg His Ala Ser Ile Ala Asn Lys Trp Lys Thr Ile His Leu Ile Met
292             580           585                   590
295 His Ser Gly Met Asn Ala Thr Thr Ile His Phe Asn Leu Thr Phe Gln
296         595           600                   605
299 Asn Thr Asn Asp Glu Glu Phe Lys Met Gln Ile Thr Val Glu Val Asp
300     610           615               620
303 Thr Arg Glu Gly Pro Lys Leu Asn Ser Thr Ala Gln Lys Gly Tyr Glu
304 625           630               635               640
307 Asn Leu Val Ser Pro Ile Thr Leu Leu Pro Glu Ala Glu Ile Leu Phe
308             645           650               655
311 Glu Asp Ile Pro Lys Glu Lys Arg Phe Pro Lys Phe Lys Arg His Asp
312         660           665               670
315 Val Asn Ser Thr Arg Arg Ala Gln Glu Glu Val Lys Ile Pro Leu Val
316         675           680               685
319 Asn Ile Ser Leu Leu Pro Lys Asp Ala Gln Leu Ser Leu Asn Thr Leu
320     690           695               700
323 Asp Leu Gln Leu Glu His Gly Asp Ile Thr Leu Lys Gly Tyr Asn Leu
324 705           710               715               720
327 Ser Lys Ser Ala Leu Leu Arg Ser Phe Leu Met Asn Ser Gln His Ala
328             725           730               735
331 Lys Ile Lys Asn Gln Ala Ile Ile Thr Asp Glu Thr Asn Asp Ser Leu
332             740           745               750
335 Val Ala Pro Gln Glu Lys Gln Val His Lys Ser Ile Leu Pro Asn Ser
336         755           760               765
339 Leu Gly Val Ser Glu Arg Leu Gln Arg Leu Thr Phe Pro Ala Val Ser
340     770           775               780
343 Val Lys Val Asn Gly His Asp Gln Gly Gln Asn Pro Pro Leu Asp Leu
344 785           790               795               800
347 Glu Thr Thr Ala Arg Phe Arg Val Glu Thr His Thr Gln Lys Thr Ile
348             805           810               815
351 Gly Gly Asn Val Thr Lys Glu Lys Pro Pro Ser Leu Ile Val Pro Leu
352         820           825               830
355 Glu Ser Gln Met Thr Lys Glu Lys Lys Ile Thr Gly Lys Glu Lys Glu
356         835           840               845
359 Asn Ser Arg Met Glu Glu Asn Ala Glu Asn His Ile Gly Val Thr Glu
360     850           855               860
363 Val Leu Leu Gly Arg Lys Leu Gln His Tyr Thr Asp Ser Tyr Leu Gly
364 865           870               875               880
367 Phe Leu Pro Trp Glu Lys Lys Lys Tyr Phe Leu Asp Leu Leu Asp Glu
368             885           890               895
371 Glu Glu Ser Leu Lys Thr Gln Leu Ala Tyr Phe Thr Asp Ser Lys Asn
372         900           905               910
375 Arg Ala Arg Tyr Lys Arg Asp Thr Phe Ala Asp Ser Leu Arg Tyr Val
376         915           920               925
379 Asn Lys Ile Leu Asn Ser Lys Phe Gly Phe Thr Ser Arg Lys Val Pro
```

```
380      930                 935                 940
383 Ala His Met Pro His Met Ile Asp Arg Ile Val Met Gln Glu Leu Gln
384 945                 950                 955                 960
387 Asp Met Phe Pro Glu Glu Phe Asp Lys Thr Ser Phe His Lys Val Arg
388                 965                 970                 975
391 His Ser Glu Asp Met Gln Phe Ala Phe Ser Tyr Phe Tyr Tyr Leu Met
392             980                 985                 990
395 Ser Ala Val Gln Pro Leu Asn Ile  Ser Gln Val Phe Asp  Glu Val Asp
396             995                 1000                1005
399 Thr Asp  Gln Ser Gly Val Leu  Ser Asp Arg Glu Ile  Arg Thr Leu
400      1010                 1015                1020
403 Ala Thr  Arg Ile His Glu Leu  Pro Leu Ser Leu Gln  Asp Leu Thr
404      1025                 1030                1035
407 Gly Leu  Glu His Met Leu Ile  Asn Cys Ser Lys Met  Leu Pro Ala
408      1040                 1045                1050
411 Asp Ile  Thr Gln Leu Asn Asn  Ile Pro Pro Thr Gln  Glu Ser Tyr
412      1055                 1060                1065
415 Tyr Asp  Pro Asn Leu Pro Pro  Val Thr Lys Ser Leu  Val Thr Asn
416      1070                 1075                1080
419 Cys Lys  Pro Val Thr Asp Lys  Ile His Lys Ala Tyr  Lys Asp Lys
420      1085                 1090                1095
423 Asn Lys  Tyr Arg Phe Glu Ile  Met Gly Glu Glu Glu  Ile Ala Phe
424      1100                 1105                1110
427 Lys Met  Ile Arg Thr Asn Val  Ser His Val Val Gly  Gln Leu Asp
428      1115                 1120                1125
431 Asp Ile  Arg Lys Asn Pro Arg  Lys Phe Val Cys Leu  Asn Asp Asn
432      1130                 1135                1140
435 Ile Asp  His Asn His Lys Asp  Ala Gln Thr Val Lys  Ala Val Leu
436      1145                 1150                1155
439 Arg Asp  Phe Tyr Glu Ser Met  Phe Pro Ile Pro Ser  Gln Phe Glu
440      1160                 1165                1170
443 Leu Pro  Arg Glu Tyr Arg Asn  Arg Phe Leu His Met  His Glu Leu
444      1175                 1180                1185
447 Gln Glu  Trp Arg Ala Tyr Arg  Asp Lys Leu Lys
448      1190                 1195
451 <210> SEQ ID NO: 3
452 <211> LENGTH: 5597
453 <212> TYPE: DNA
454 <213> ORGANISM: Homo sapiens
456 <400> SEQUENCE: 3
457 cggagccgag cgggcgtccg tcgccggagc tgcaatgagc ggcgcccgga ggctgtgacc      60
459 tgcgcgcggc ggcccgaccg gggcccctga atggcggctc gctgaggcgg cggcggcggc     120
461 ggcggctcag gctcctcggg gcgtggcgtg gcggtgaagg ggtgatgctg ttcaagctcc     180
463 tgcagagaca aacctatacc tgcctgtccc acaggtatgg gctctacgtg tgcttcttgg     240
465 gcgtcgttgt caccatcgtc tccgccttcc agttcggaga ggtggttctg gaatggagcc     300
467 gagatcaata ccatgttttg tttgattcct atagagacaa tattgctgga aagtcctttc     360
469 agaatcggct ttgtctgccc atgccgattg acgttgttta cacctgggtg aatggcacag     420
471 atcttgaact actgaaggaa ctacagcagg tcagagaaca gatggaggag gagcagaaag     480
473 caatgagaga aatccttggg aaaaacacaa cggaacctac taagaagagt gagaagcagt     540
```

Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding
explanation is presented in the <220> to <223> fields of
each sequence using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/023,888

DATE: 01/15/2002
TIME: 19:00:26

Input Set : A:\203515US77.txt
Output Set: N:\CRF3\01152002\J023888.raw

```
L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1639 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1925 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:1929 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:1931 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
```